Why the energy efficiency gap is smaller than we think: quantifying heterogeneity and persistence in the returns to energy efficiency measures

Daire McCoy[1,3]*, Raphaela Kotsch[1,2]

[1] *Grantham Research Institute, London School of Economics, UK*

[2] *The Department of Geography and Environment, London School of Economics, UK*

[3] *The Economic and Social Research Institute, Ireland*

September 3, 2018

PRELIMINARY DRAFT - COMMENTS WELCOME - PLEASE DO NOT CITE

Abstract

The "energy-efficiency gap" is a topic that has received much attention in the academic literature. While the role of market and behavioural failures have been discussed at length, much less focus has been on quantifying the magnitude of heterogeneity and persistence that exists in the realised savings from installing measures. This paper systematically explores variation in the returns to energy efficiency upgrades. Statistical matching and panel econometric estimations are employed on a database of over four million households over an eight year period to mitigate selection bias into various government schemes which funded the upgrades and to control for unobserved heterogeneity. Detailed characterisation of both cross-sectional and temporal variation in the energy savings associated with a number of widely used energy efficiency measures is presented. This allows an assessment of the persistence of savings over longer periods of time than is typically examined, and an examination of the distributional impacts. The econometric estimates are then combined with cost-estimates and a range of future energy-price scenarios to determine the cost-effectiveness of measures. The results raise concern over the distributional effects of energy efficiency measures and policies. Not only do households in more deprived areas experience lower energy savings, the savings erode more quickly over time for these households. This result has important implications for improving our

*Corresponding author. *E-mail address*: d.m.mccoy@lse.ac.uk

understanding of the investment incentives households face and also for improving our evaluations of energy efficiency policies. It also suggests that the energy-efficiency gap requires less explanation than some would suggest.

Keywords: Energy efficiency gap; policy evaluation; panel data econometrics

JEL codes: C23; D12; Q40; Q48

1 Introduction

The European Union's 2020 strategy, which constitutes a set of binding legislation, aims to cut greenhouse gas (GHG) emissions by 20% by 2020 compared to 1990 levels. Reducing energy demand plays a crucial role in reaching this goal. With the *2008 Climate Change Act*, the United Kingdom (UK) committed to the ambitious target of an 80% reduction in GHG emissions by 2050 relative to 1990 levels. The UK's domestic sector is responsible for around 25% of GHG emissions and accounts for around 30% of total final energy consumption, mainly from gas and electricity consumption[1].

A key element in reducing domestic energy consumption, is encouraging consumers to install energy efficiency measures, through either policy or market-based instruments. The reluctance of some consumers to make energy saving investments that offer them seemingly positive net-present value (NPV) returns has been widely studied[2]. A recent paper by Gerarden et al. (2015) characterises this problem into three district components: market-failures, behavioural explanations and model and measurement errors. The authors argue that the energy efficiency gap may not be as large as expected and that unobserved costs, overstated savings from adoption, consumer heterogeneity, inappropriate discount rates and uncertainty may all contribute to the low adoption rate not being as "*paradoxical as it first appears.*". Model and measurement error is also a factor that affects policy evaluations, many of which rely on ex-ante engineering estimates of savings, or do not take behavioural responses into account.

Recent research has highlighted the difference between engineering estimates of energy savings and actual realised savings, finding that engineering estimates can overstate the actual savings by as much as 2.5 times (Fowlie et al., 2015). Even in ex-ante analyses which use observed rather than modelled data, specific factors related to usage patterns in any particular period may bias results both before and after, while poor installation quality or degradation in the installed equipment may affect the results post-installation. Variation over time could affect the accuracy of measurement, the attractiveness of the investment, or the cost-effectiveness of a government scheme. Further, variations in energy prices both before and after the installation may affect both expectations and realisations of the investment's net-present value.

Time-scale has proven an important factor when examining the impact of building energy codes on energy consumption (Kotchen, 2017). However, most evaluations of energy efficiency improvements take a short time-

[1]48% and 41% respectively (Parag and Darby, 2009)

[2]For example see Hausman (1979); Blumstein et al. (1980); Jaffe and Stavins (1994); Golove and Eto (1996); Allcott and Greenstone (2012).

scale, usually a window of 1-2 years on either side of the intervention (Adan and Fuerst, 2015; Fowlie et al., 2015; Hamilton et al., 2016).

This research contributes by providing information on the extent of heterogeneity that exists with regard to the savings associated with installing different energy efficiency measures. Uniquely, we also demonstrate how the savings from measures change over time for different household types. Not only do households in more deprived areas experience lower energy savings, the savings erode more quickly over time - in some cases the savings reduce by 50 percent within six years. This result has important implications for improving our understanding of the investment incentives households face and also for improving our evaluations of energy efficiency policies.

In order to conduct this analysis we exploit an extremely large database of home energy efficiency upgrades and metered energy consumption[3], covering over four million households and a period of eight years. By combining statistical matching and a range of panel econometric estimators we control for unobserved heterogeneity and selection into various government schemes which funded the upgrades. Our database covers the universe of households entering into energy efficiency schemes administered by energy suppliers in the UK, thus reducing the potential for "site-selection bias" as identified by Allcott (2015).

The data allows us to examine the variation in performance depending on when measures were installed; how they perform over time; how this varies by dwelling and socioeconomic characteristics; and ultimately how this affects the cost-effectiveness of measures for different household types. Results indicate significant cross-sectional and temporal variation in energy savings; that the persistence of savings varies by the type of measure installed and the socioeconomic characteristics of the household. The measures are generally still NPV positive under a range of price scenarios, but the returns are much lower than expected. This research also raises concerns over distributional factors given how the costs of policies are subsequently levied on households.

The rest of the paper is organised as follows; Section 2 provides the context in which this analysis takes place; Section 3 the data; Section 4 describes the methodological approach employed and considerations undertaken; Section 5 outlines the results; Section 6 provides a concluding discussion.

[3]The National Energy Efficiency Framework Database (NEED). Further details available at: https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework

2 Background

The Supplier Obligation (SO), first introduced to the UK in 1994, has become the principle policy instrument for implementing energy efficiency improvements in the domestic sector in the UK (Rosenow, 2012). The Supplier Obligations are an example of a “Tradable White-Certificate” (TWC) scheme. These are regulatory mechanisms, employing a market-based approach to deliver energy savings. Theoretically they can be considered a hybrid subsidy-tax instrument, in which suppliers provide subsidies for energy efficiency upgrades that are then recovered through increased energy prices (Giraudet et al., 2012), having parallels with traditional demand-side management (DSM) programmes in that companies are required to invest in projects that ultimately reduce demand for their product (Sorrell et al., 2009).

As outlined in Bertoldi and Rezessy (2008) and Giraudet et al. (2012), SOs have three main features: an obligation is placed on energy companies to achieve a quantified target of energy savings, savings are based on standardised ex-ante calculations, the obligations can be traded with other obligated parties. This flexibility ideally allows suppliers to choose the most cost-effective way to reach their target. Suppliers bear the cost of installations in the first instance, costs are then passed through to their entire population of customers through increases in energy prices (Chawla et al., 2013). Clearly, this may have distributional consequences if certain segments of the population are less likely to avail of the schemes - to alleviate this concern, targets were imposed regarding the proportion of savings to be achieved from lower income groups.

The Department of Energy and Climate Change (DECC) [now Department for Business, Energy and Industrial Strategy (BEIS)], sets the savings targets which are then enforced by the energy regulator, the Office of Gas and Electricity Markets (Ofgem). Ofgem sets and administers individual savings targets for each energy supplier. Energy suppliers have various options to achieve their targets such as contracting installers, subsidising energy efficiency products, cooperating with local authorities, delivery agents or supermarkets, or directly working with their customers (Rosenow, 2012).

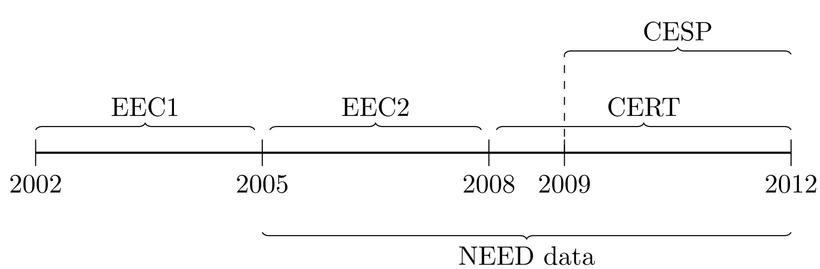

Figure 1: UK Energy Efficiency Programmes 2005-2012

Figure 1 gives an overview of SOs from 2002-2012. The first Energy Efficiency Commitment (EEC1) ran from 2002 to 2005, followed by EEC2 in 2005. In 2008, EEC2 was replaced by the Carbon Emissions Reduction Target (CERT) which ran until 2012. In 2009, the Community Energy Saving Programme (CESP) was introduced in parallel with CERT. While the main architecture of SOs did not change, the savings targets and the costs of the delivering the programmes increased over time. Rosenow (2012) provides a comprehensive overview of the main changes in each scheme from 1994 - 2012 with regards to the target, the costs, social equity implications and other changes in design. The main change concerned the target size, increasing substantially in lifetime savings from 2.7 to 494 terawatt hours (TWh) between 1994 and 2012 [4] (Rosenow, 2012).

From 2002, all programmes included a target for disadvantaged households and fuel poverty increasingly came to the fore. Eventually, CESP only allowed projects to be carried out in specific low income areas of Britain, the lowest 10- 15% of areas ranked in Income Domain of the Indices of Multiple Deprivation (Hough and Page, 2015). Thus, CESP was only available in certain geographical regions. Furthermore, CESP introduced a new bonus structure that incentivised the installation of multiple measures in a single dwelling and the treatment of as many dwellings as possible in the same area (Duffy, 2013). Table 1 summarises the main changes for the schemes under consideration.

Table 1: Overview of Supplier Obligations

	ECC1	ECC2	CERT	CESP
Target	62 TWh	130 TWh	494 TWh (293 million t CO2)	19.25 Mt CO2
Annual costs (millions)	167	400	1,158	unknown
% savings in priority group	50%	50%	40%	lowest 10-15% of areas ranked by IMD
Number of cavity wall insulations	791,524	1,760,828	2,568,870	3,000
# Number of loft insulations	754,741	1,780,302	3,897,324	23,503
# Number of replacement heating systems	366,488	2,018,812	31,986	42,898

Source: Lees (2006, 2008); Rosenow (2012); Duffy (2013)

[4] 1 TWh is equal to $1e^9$ kWh

A key feature of all previous evaluations of the above policies is that the energy savings achieved were based on model ex-ante estimates and not actual ex-post data. As stated above model estimates tend to overstate actual savings significantly. This would lead to concern over the accuracy of measurement regarding both the energy savings achieved and the cost-effectiveness of various policies in delivering savings.

3 Data

The National Energy Efficiency Database (NEED) contains dwelling-level data on four million households, over an eight-year period. Information comes from a range of sources including meter point electricity and gas consumption data, Valuation Office Agency (VOA) property attribute data, the Homes Energy Efficiency Database (HEED) containing data on energy efficiency measures installed, and data modelled by Experian on household characteristics. Details are in Table 2.

Table 2: Data sources combined in NEED

Type of variable	Source
Energy efficiency measures	HEED/Ofgem/DECC
Energy consumption	Energy Supplier
Property attributes	VOA
Household characteristics	Experian

3.1 Measures installed

The database includes measures installed through EEC2, CERT and CESP. Unfortunately, it does not contain information on direct subsidy schemes in the UK, such as the Warm Front scheme. However, Supplier Obligations were by far the most prevalent mechanism for delivering energy savings in residential dwellings in the UK over this period. Measures included are cavity-wall insulation, loft insulation and boiler replacements. In total over two million measures were installed over the period of our analysis, this is graphically represented in Figure 2.

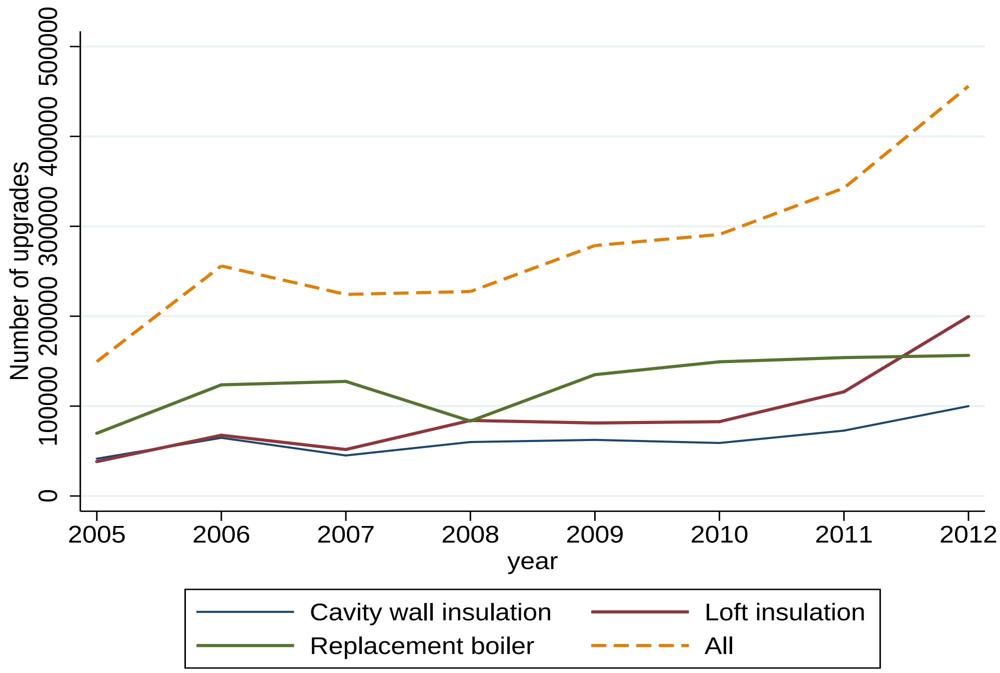

Figure 2: Energy efficiency measures installed, 2005-2012

The NEED database does not include an exhaustive list of measures installed as part of the various schemes, appliances and lighting also featured but are not included. However, as Table 3 demonstrates, insulation and heating comprised the vast majority of estimated energy savings across various schemes over this period.

Table 3: Energy savings by scheme and measure

	EEC1	EEC2	CERT
	2002-2005	2005-2008	2008-2012
Insulation	56%	75%	66.20%
Heating	9%	8%	8.20%
Lighting	24%	12%	17.30%
Appliances	11%	5%	5.90%
Other	-	-	2.40%

Source: Lees (2006, 2008); Ofgem (2013)

All insulation installations in our dataset were funded through government schemes. In the early part of our sample (pre-2007) boiler installations were also likely to have been funded through government schemes, however government support for replacement boilers was withdrawn during EEC2, as a combination of previous support schemes and new building regulations in 2005 had already delivered a significant penetration of new condensing boilers. Therefore the boiler data we report on is a combination of publicly and privately funded investments.

3.2 Energy consumption

Figure 3 illustrates that on average, gas consumption reduced by 27% between 2005 and 2012 and electricity consumption reduced by 14%. Both of these trends are encouraging signs that the various polices in place over this period were having an effect.

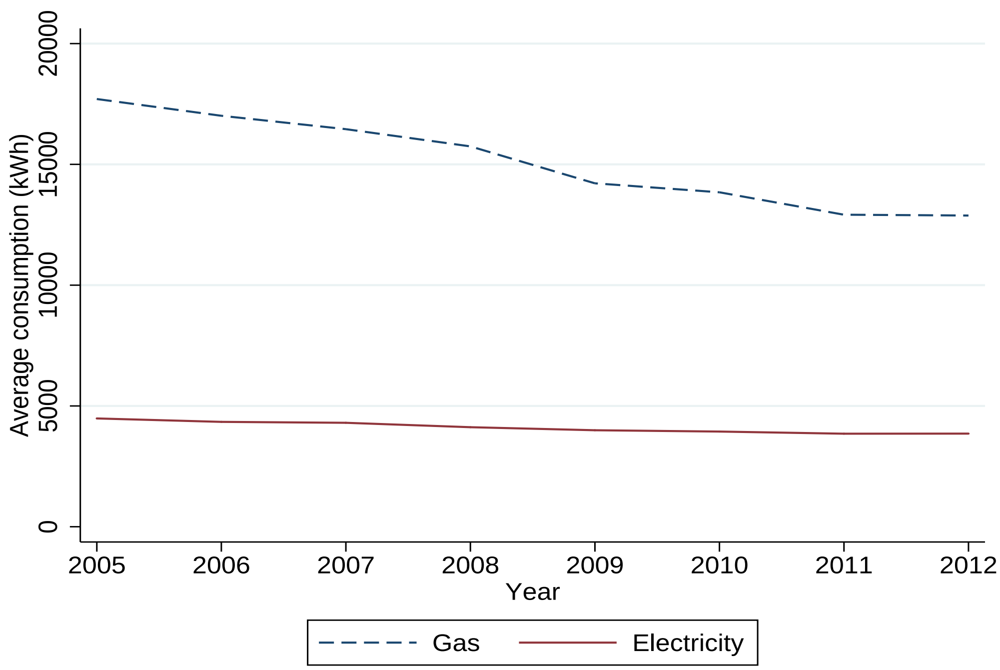

Figure 3: Average domestic energy consumption UK, 2005-2012

3.3 Socioeconomic characteristics

The NEED dataset comprises information on household characteristics modelled by Experian and matched with indicators based on the geographic location of the property (DECC, 2016). For reasons of data protection, the dataset was anonymised and household-level information on variables such as income and tenure-type are not available. However, the dataset does include two composite indicators of the socio-economic background of the households.

1. *Index of multiple deprivation* (IMD)

   NEED contains two variables describing IMDs: IMD 2010 for England and IMD 2011 for Wales. Both indicators classify Lower Layer Super Output Areas (LSOAs) according to a quintile ranking that is based on eight different domains that are incorporated using a weighting scheme. The first quintile (IMD=1) indicates the most deprived areas. Table 4 shows the composition of domains that are incorporated in the indicators and their weight in percent (Payne and Abel, 2012; of National Statistics, 2011).

Table 4: Composition of IMD in %

	England 2010	Wales 2011
Income	22.5	23.5
Employment	22.5	23.5
Health	13.5	14
Education	13.5	14
Access/barriers to services	9.3	10
Living environment/ housing	9.3	5
Physical environment	0	5
Crime [Wales: Community Safety]	9.3	5

2. *Fuel poverty indicator* (FP)

   Combining data from the English Housing Survey and Census data, the fuel poverty indicator indicates if households are fuel poor based on the households' income and energy requirements, as well as on fuel prices (BEIS, 2013).

4 Econometric approach

4.1 The model

We are interested in assessing the extent to which energy efficiency upgrades affect energy consumption. Energy consumption is determined by a range of factors such as temperature, characteristics of the dwelling and its inhabitants, and energy prices. The following baseline specification is estimated:

$$ln(y_{it}) = \alpha_i + \gamma_t + \rho_{rt} + \delta \sum_{j=1}^{3} D_{ijt} + \epsilon_{it} \tag{1}$$

Where y_{it} denotes energy consumption by household i in year t, α_i is a household fixed-effect, γ_t is a year fixed-effect which controls for unobserved factors which vary at an annual level such as broader macroeconomic conditions and weather patterns, ρ_{rt} is a year-by-region fixed effect to control for factors which vary at a sub-national level, such as more localised economic shocks and weather patterns, D_{it} is the treatment dummy. The key parameter of interest is δ the average treatment effect on the treated (ATT).

The model is estimated as a first-differenced fixed effects panel specification controlling for unobserved time-invariant household characteristics which might affect energy consumption. Year fixed effects control for annual trends which may affect all households at different points in time, such as economic shocks or extreme weather, year by region fixed effects control for annual trends that might vary at a more dissagregate level, localised economic shocks or weather patterns will be absorbed by this parameter.

In the course of the analysis, a variety of extensions to the above are estimated, to account for interactions between different upgrades and to examine the performance of upgrades over time. The following specification captures interactions between different upgrades:

$$ln(y_{it}) = \alpha_i + \gamma_t + \rho_{rt} + +\lambda_t + W_{it}\beta + \delta_1 \sum_{j=1}^{3} D_{ijt} + \delta_2 \sum_{j=1}^{3} D_{ijt}. \sum_{j=1}^{3} D_{ijt} + \epsilon_{it} \tag{2}$$

Following this we examine the effect of upgrades over time. Year one is used as a control period and the ATT for all subsequent periods is estimated. The below specification is estimated:

$$ln(y_{it}) = \alpha_i + \sum_{t=2}^{7} \gamma_t + \rho_{rt} + \lambda_t + W_{it}\beta + \delta_1 \sum_{j=1}^{3} D_{ijt} + \delta_2 \sum_{j=1}^{3} D_{ijt}. \sum_{t=2}^{7} \gamma_t + \epsilon_{it} \tag{3}$$

All models are estimated for both gas and electricity consumption. Standard errors are clustered at the household level in all specifications. The data allow us to create multiple treatment and control groups. Treatment groups are created for the entire sample period and for each individual year of upgrade. This allows us to examine how treatment effects vary over time.

4.2 Identification

4.2.1 The problem of unobserved heterogeneity

The fixed effects estimators described above are based on the assumption of conditional mean independence or unconfoundedness, selection on observables or ignorability (Caliendo and Kopeinig, 2005; Angrist and Pischke, 2009; Wooldridge, 2010), which requires that both of the following equations hold:

$$E[Y_{it}^0|A_i, t, X_{it}, D_{it}] = E[Y_{it}^0|A_i, t, X_{it}] \tag{4}$$

and

$$E[Y_{it}^1|A_i, t, X_{it}, D_{it}] = E[Y_{it}^1|A_i, t, X_{it}] \tag{5}$$

Thus, it assumes that D_{it} is strictly exogenous and as good as randomly assigned conditional on A_i (Angrist and Pischke, 2009). Furthermore, both fixed effects, the time effects γ and household effects α are assumed to be additive and homogenous (Ferraro and Miranda, 2017):

$$E[Y_{it}^1|A_i, t, X_{it}] = E[Y_{it}^0|A_i, t, X_{it}] + \delta \tag{6}$$

As we are primarily interested in the effect on the households who availed of the schemes - the average treatment effect on the treated (ATT), and not necessarily the effect on the whole population - the average treatment effect (ATE), the condition of unconfoundedness can be relaxed and equation (5) can be ignored.

There is strong evidence that the presence of unobserved heterogeneity leads to inaccurate estimates of the ATE and ATT in a fixed-effects OLS setting (Ferraro and Miranda, 2017; Gibbons et al., 2014). Self-selection bias occurs as households voluntarily decide to apply upgrades in their homes or take part in government funded schemes, potentially causing the treatment and control group to differ systematically in aspects that both affect their likelihood of taking part in energy efficiency programs, and their energy consumption, causing the failure of the conditional mean independence assumption (Wooldridge, 2010). Unobserved heterogeneity between households means that households respond differently to common shocks. For instance, increasing energy prices might lead to different behaviour of low and high income households. Second, the crucial assumption of a linear model with additive and homogeneous effects implies that the fixed effect estimates give a weighted average based on the frequency of groups as well as the sample variances within groups (Gibbons et al., 2014). This is problematic as the fixed effects estimator overweights groups that have larger variance of treatment conditional upon other covariates and underweights groups with smaller conditional variance if heterogenous treatment is prevalent (Ferraro and Miranda, 2017). One strategy to overcome this threat and to obtain consistent and unbiased estimators is to pre-process the data through statistical matching (Wooldridge, 2010). The following section outlines this approach.

4.3 Matching

Policy evaluations of secondary data typically employ statistical matching, along with differences-in-differences estimation, or exploit the longitudinal nature of the data with a panel fixed-effects specification. However, both of these measures may suffer from bias through either unobserved temporal effects or unobserved heterogeneity. Recent research has shown that by combining these methodologies, the accuracy of evaluations can approach that achieved by a randomised-controlled trial (RCT) (Ferraro and Miranda, 2017).

Coarsened-exact matching (CEM) is a non-parametric statistical procedure which improves the estimation of causal effects by reducing imbalance in observed variables between treatment and control groups (Iacus et al., 2008; Blackwell et al., 2009). In this case we are concerned with balancing the group that received the energy efficiency upgrades with the group that did not. By balancing all observed variables we can isolate the effect of the upgrade on energy consumption. Alberini and Towe (2015) use a similar approach in an analysis of analysis of home energy audits in the state of Maryland.

4.3.1 Justification for choice of matching covariates

Covariates on which the matching is performed should be predictors of household energy consumption and simultaneously impact the uptake of energy efficiency upgrades. The IMD of the area in which the household resides, provides important information on the household's socioeconomic environment, an important predictor of energy consumption and energy efficiency uptakes. Hamilton et al. (2014) finds a strong relationship between the uptake rate of energy efficiency upgrades and neighbourhood income levels.

While more specific information on household socioeconomic characteristics, such as employment status, income and health are significant predictors of energy expenditures, they are found to have a smaller impact than that of dwelling characteristics and household size (Longhi, 2015). The period in which the dwelling was built has an important impact on residential energy consumption (Brounen et al., 2012; Harold et al., 2015). In order to account for regional differences in weather patterns, we include a variable reflecting the region in which the dwelling is located. Alberini and Towe (2015) provide evidence that matching solely based on dwelling or household characteristics is not sufficient and can be optimised if past energy usage is also included. By performing matching on energy consumption in prior years, we can account for unobservable household and property characteristics that might vary over time, such as the household size, composition and appliance usage. Taking into consideration all these factors, matching is performed on the following variables: property age, fuel-type, energy consumption in prior years, region and the IMD of the area in which the household resides.

4.3.2 Quality of matching

The quality of the matching process depends on the similarity in the distribution of covariates between treated and matched control group. This is commonly assessed by comparing the standardised difference and variance ratio of the variables in both groups, before and after matching (Caliendo and Kopeinig, 2005). The standardised difference is the difference in sample means in the treated and control group, divided by the corresponding sample variances. Formally:

$$d = \frac{\overline{\mathrm{x}}_{treatment} - \overline{\mathrm{x}}_{control}}{\sqrt{\frac{s^2_{treatment} + s^2_{control}}{2}}} \tag{7}$$

It allows for a comparison of balance which is independent of the sample size and measurement unit (Austin, 2009). The smaller the difference, the better, and it is recommended that this ratio should not exceed 10 percent (Austin, 2009).

The variance ratio measures the ratio of the mean variance in the treated and control group for each covariate. Formally:

$$F = \frac{s^2_{treatment}}{s^2_{control}} \tag{8}$$

This should be close to unity (Austin, 2009; Ferraro and Miranda, 2017). A significant divergence from this indicates that the matching model is misspecified. Further methods of balance diagnostic include assessing the magnitude of the difference between treatment and matched control group covariates using tests for statistical significance. However, the use of the t-test for balance testing is criticised for several reasons under which the most problematic is the dependence on the sample size. For instance, randomly discarding control units will always increase the balance, falsely indicating a better balance (Imai et al., 2008).

As can be demonstrated by Figure 4 and Tables A1 and A2 our extremely large sample size allows a high level of precision in matching. A high degree of balance is achieved on both variables used in matching and variables not used in matching as can be seen from the standardised differences, variance ratios and the distributions of matched electricity and gas consumption.

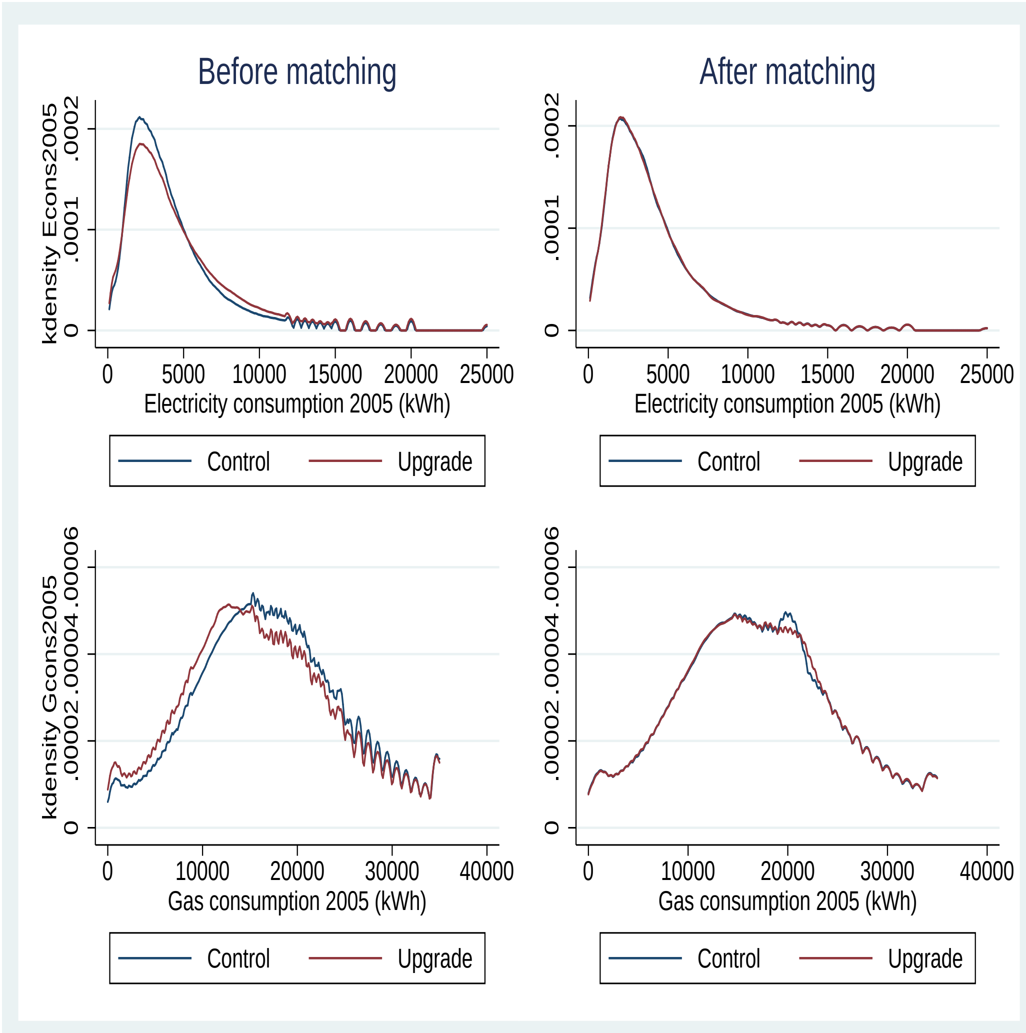

Figure 4: Energy consumption before and after matching

Another important element in assessing the quality of matching is that the parallel paths assumption is not violated. This assumption states that without treatment, the average change for the treated would have been equal to the observed average change in the controls. Figure 5 demonstrates that this assumption holds for all treatment and control groups.

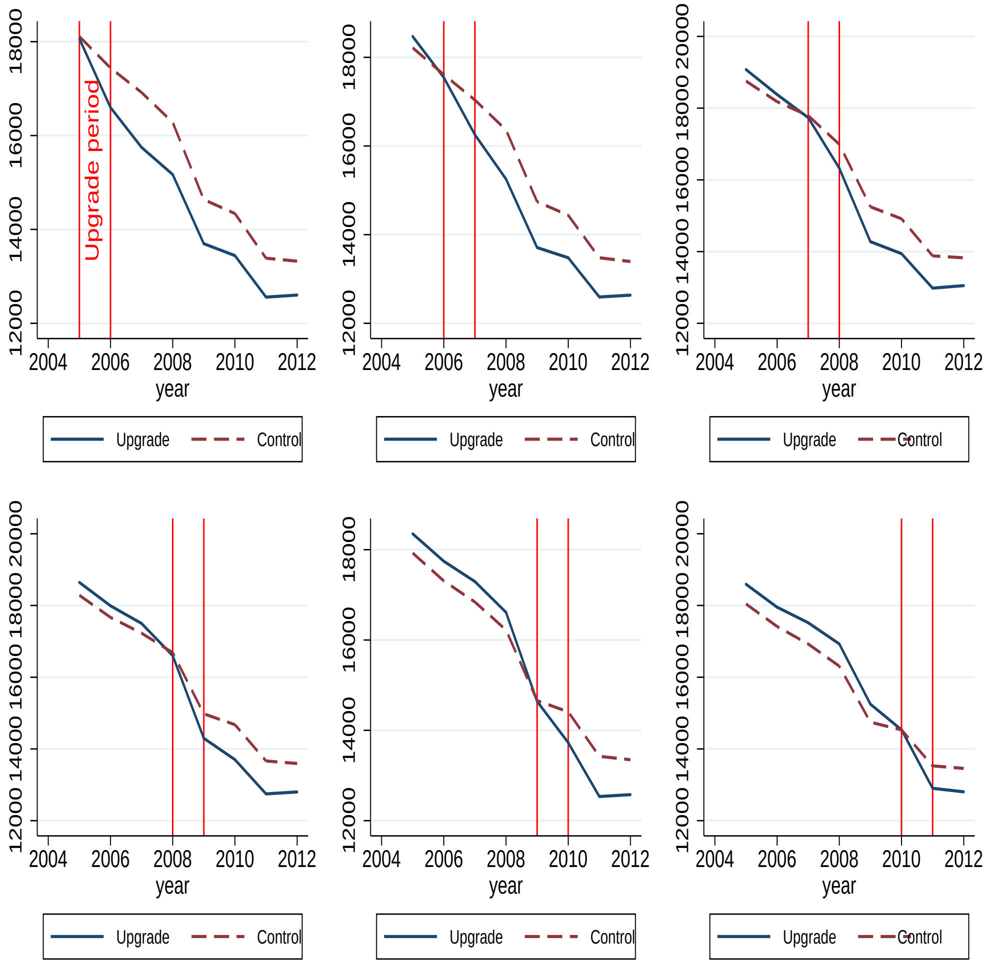

Figure 5: Energy consumption trend in upgrade and control group

5 Results

All reported results are estimates of the average treatment effect on the treated (ATT) and can be interpreted as percentage energy savings. Multiple upgrade and control groups are created (see the notes below Table **??** for example) for the entire period of analysis and for each individual year. This allows us to calculate the average effect and to examine trends over time. Analysis is restricted to households with electricity consumption between 100 and 25,000 kWh, and gas consumption between 3000 and 50,000 kWh. Outliers are excluded to minimise risk of inclusion of invalid consumption readings or non-domestic properties. Following this we create dummy variables to indicate if household energy (either electricity or gas) changed by more than 50, 60 or 70 percent in any given year. These dummy variables are then used in sensitivity analysis to control for unobserved changes in occupancy that might cause large changes in consumption.

5.1 The effect of energy efficiency upgrades by year of upgrade

Table 5 shows that the energy savings over time are quite consistent for each measure, regardless of when the installation took place. Annual gas savings for cavity wall insulation range from 8-11 percent, loft insulation 2-3 percent, and replacement heating systems 8-10 percent.

Table 5: The effect of energy efficiency upgrades on energy consumption

	(1) Full sample	(2) 2006 upgrades	(3) 2007 upgrades	(4) 2008 upgrades	(5) 2009 upgrades	(6) 2010 upgrades	(7) 2011 upgrades
Cavity wall insulation	-0.094***	-0.097***	-0.111***	-0.099***	-0.098***	-0.097***	-0.101***
	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Loft insulation	-0.030***	-0.026***	-0.031***	-0.028***	-0.027***	-0.039***	-0.035***
	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Replacement boiler	-0.092***	-0.080***	-0.093***	-0.087***	-0.102***	-0.109***	-0.099***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Control variables	Y	Y	Y	Y	Y	Y	Y
Household fixed effects	Y	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y	Y
Year*region fixed effects	Y	Y	Y	Y	Y	Y	Y
Observations		617022	545627	564756	730447	746573	871379
Number of households		77128	68203	70595	91306	93322	108922
R squared	0.349	0.327	0.353	0.370	0.369	0.386	0.367

Notes: This table reports coefficient estimates and standard errors from eight separate regressions. The dependent variable in all regressions is the logarithm of annual gas consumption in kilowatt hours. Column(1) "All" denotes efficiency upgrades occurring at any time during the sample period. Columns (2-8) relate to upgrades occurring only in the relevant year. Each individual year denotes upgrades occurring solely in that year. For each upgrade group a matched control group is created using coarsened-exact matching. The sample includes billing records from 2005 to 2012. Standard errors are clustered at the household level. Triple asterisks denote statistical significance at the 1% level; double asterisks at the 1% level; single asterisks at the 10% level.

5.2 Heterogeneity and persistence in returns to energy efficiency upgrades

5.2.1 By measure and IMD group

The next set of results, presented in Figure 6, show the interaction of the treatment variable with the variable indicating the socioeconomic characteristics of the area in which the household resides. Energy savings are much greater for those households living in more affluent areas (IMD = 5), compared to those in lower income areas (IMD = 1). This is true for all upgrades. Combining all measures, the annual savings range from approximately 15 percent for those in the lowest IMD category to approximately 25 percent for those in the highest. This result raises concerns over distributional issues as the costs of these policies were applied as a flat-rate tariff on energy bills (Chawla et al., 2013). Given that this charge is already regressive, disproportionately affecting lower income groups, to have savings concentrated in the higher income groups suggests a further loading of policy costs onto those least able to afford it.

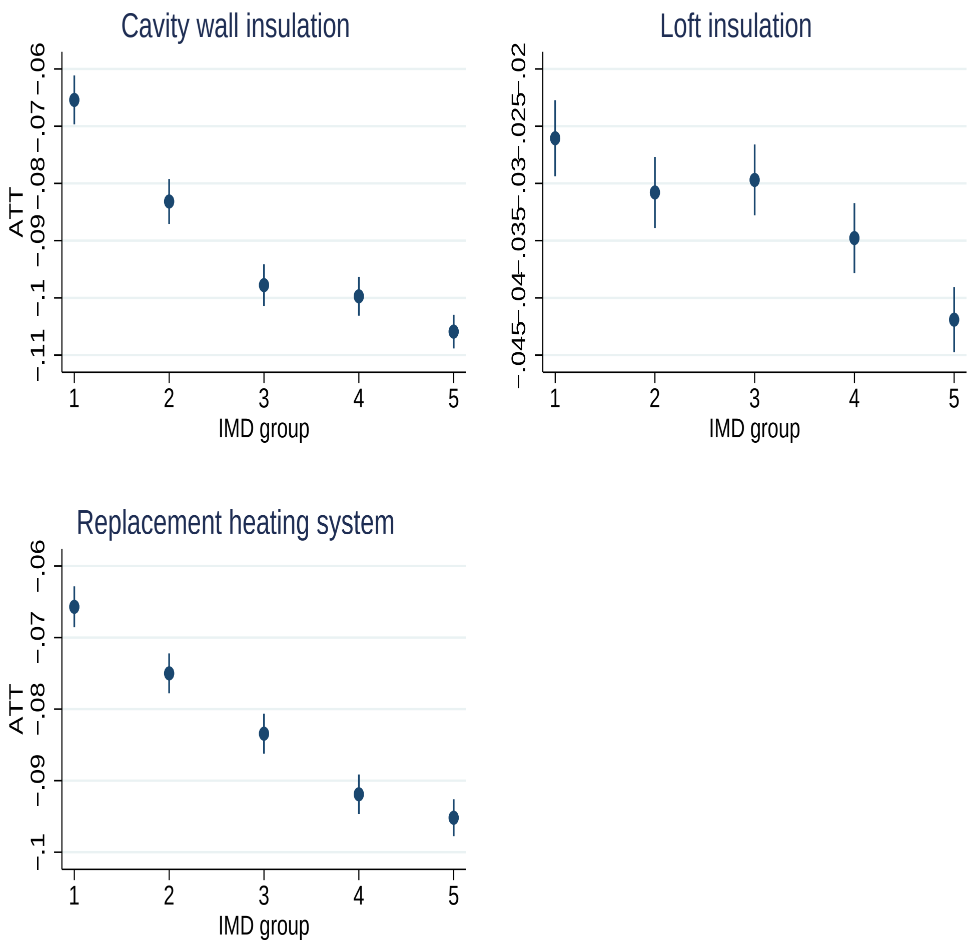

Figure 6: ATT for different IMD groups

5.2.2 By measure and over time

Figure 7 presents results of estimations in which the treatment variable is interacted with the year variable to examine the persistence of savings over time. The reported results are for upgrades occurring between 2006 and 2007. This period is chosen as it allows a matched control group to be created using 2005 consumption level, and the longest possible post-upgrade time series. Cavity wall and loft insulation show no clear time trend or degradation. This is not surprising as these measures are expected to last for 30-40 years (Dowson et al., 2012). However, for replacement heating systems the ATT shows a clear decreasing time path. This indicates that energy savings are greatest in the years immediately following installation. Given that the estimated lifespan for condensing gas boilers was 12 years around this time (Dowson et al., 2012), our results have implications for assessments of both household investment decisions and policy evaluations.

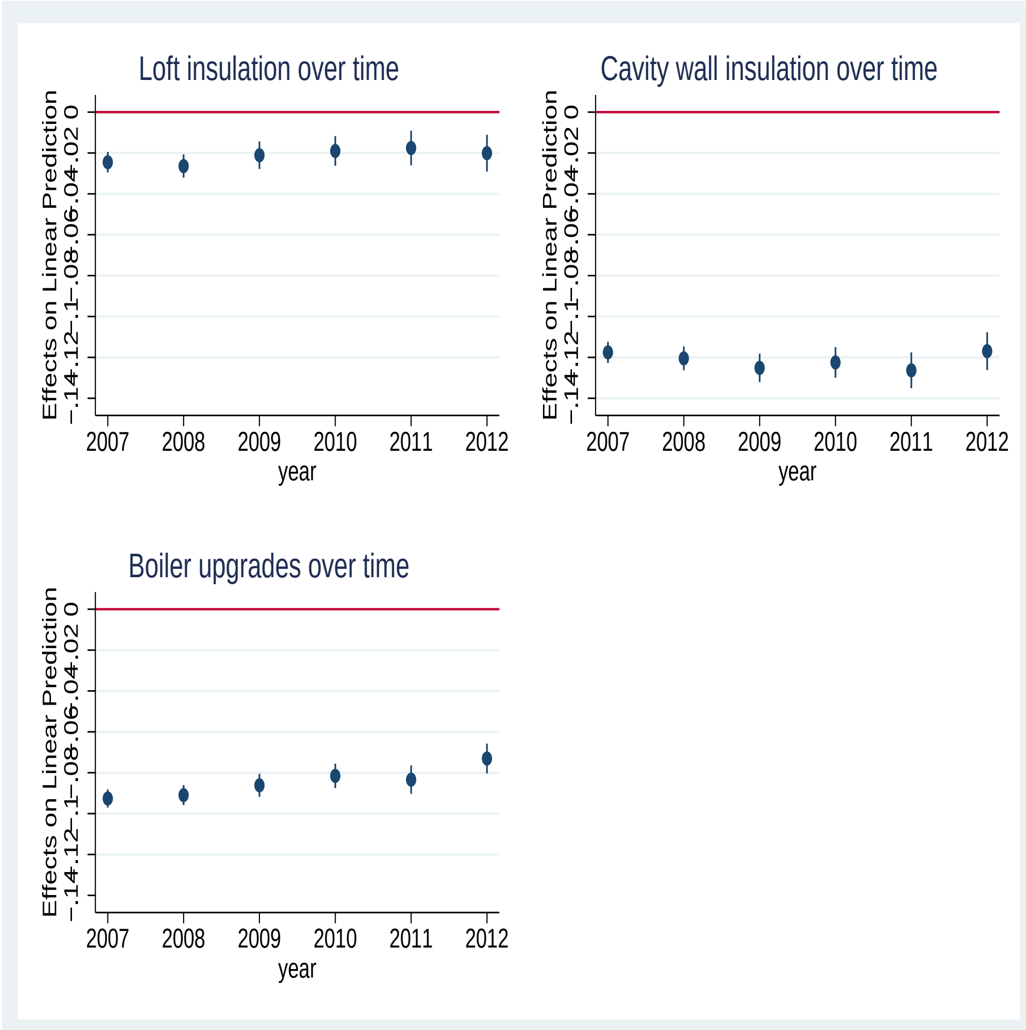

Figure 7: Persistence of ATT over time by measure

5.2.3 Heating system replacements by IMD group over time

The reduction in savings could be due to degradation of equipment or a behavioural response from households. Changes to building regulations in 2005 (needs ref) mandated all replacement boilers to be condensing and a minimum of 86% efficiency. Other than this we do not have any detailed product characteristics. However, by decomposing this trend by socioeconomic group it is possible to examine whether this effect varies for different household types. Not only are energy savings less for those in lower income areas, the trend of decreasing savings over time is much more pronounced for these households. The savings for those is the lowest IMD group have halved within five years.

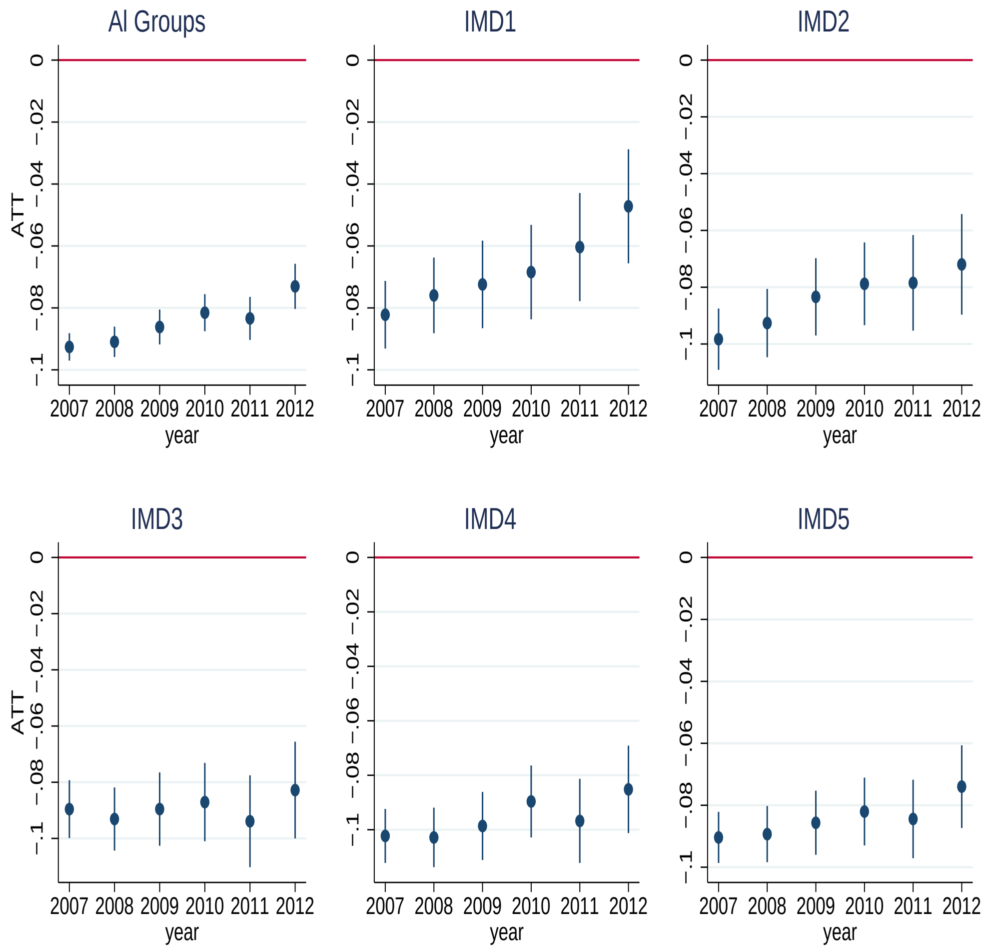

Figure 8: ATT for replacement heating system. Persistence over time and by IMD group

5.3 Cost effectiveness of measures

5.3.1 Estimated costs to suppliers and private costs of measures

In order to assess the cost-effectiveness, it is first necessary to estimate the costs. A wide degree of variation clearly exists here, and some assumptions need to be made. However, a number of published academic and policy papers provide cost-estimates. The estimates we present in Table H1 are simple averages of a range of estimates, outlined in more detail in Tables E1, E2 and E3. The costs we present are the costs incurred by the energy companies in installing each measure. These may understate the actual costs in some cases, and can thus be considered a lower bound. For example, the cost for replacement heating system used is the assumed additional cost of installing a high efficiency system, over and above a typical lower-efficiency system. The

Energy Savings Trust estimate the costs of boiler replacements to range from £700 to £6,000. On average the installation of condensing boiler costs around £2,400 per dwelling. Frontier Economics assumes that the fixed cost of for gas-fired condensing boilers lie between £2,200-3,000. (Need ref). For the purposes of comparison this is considered to be an upper bound.

Table 6: Cost assumptions for each measure

Measure	Cost assumptions (GBP)
Cavity wall insulation	350
Loft insulation	285
Replacement boiler lower (policy cost)	200
Replacement boiler upper (private cost)	2000

Source: Based on Lees (2005, 2008), Shorrock (2005)

5.3.2 Internal rate of return (IRR)

The internal rate of return (IRR) on a project is the discount rate (r) that yields a net-present value of zero, or the discount rate at which the average value of avoided discounted future energy costs equals the upfront investment cost. Formally, this can be calculated using the below formula:

$$NPV = \sum_{t=1}^{T} \frac{C_t}{(1+r)^t} - C_0 \tag{9}$$

Where T is the estimated lifespan of the measure, C_t are the avoided energy costs in year t, C_0 is the upfront investment cost and r is the IRR which we solve for. The IRR is calculated based on the econometric estimates we observe, for varying estimated lifespans of measures and assuming constant future energy prices. While these measures were largely funded by the energy companies, it is useful to estimate the private returns as this would be a critical factor in determining uptake in the absence of such schemes.

Table 7 presents estimated IRRs for each measure, calculating the IRR for 10, 20 and 30 years. Our preferred estimated lifespan is 10 years for replacement heating system and 30 years for both types of insulation. Assuming a lifespan of 30 years or more, cavity wall insulation is an attractive investment yielding a return of 16 percent. Loft insulation is less attractive yielding 6 percent. Whether to invest in a heating system depends greatly on whether one uses the estimated policy cost or private cost. The low and highly negative returns in some cases suggest that there may not be much of an energy efficiency gap to explain with regard to replacement heating systems.

Table 7: IRR for each measure

	Cavity wall	Loft	BoilerL (policy cost)	BoilerU (private cost)
IRR 10	7%	-8%	17%	-24%
IRR 20	15%	4%	23%	-7%
IRR 30	16%	6%	23%	-2%

The next set of results, presented in Table 8 presents IRR estimates for each measure and each IMD group, along with the sample average for comparison purposes. A considerable degree of variation exists around the sample average, with households living in more deprived areas experiencing much lower returns than those in more affluent areas.

Table 8: IRR for each measure and IMD group

	Sample Average	IMD1	IMD2	IMD3	IMD4	IMD5
Cavity wall	16%	11%	14%	16%	17%	18%
Loft	6%	5%	5%	5%	7%	9%
BoilerL	17%	12%	14%	17%	20%	22%
BoilerU	-24%	-26%	-25%	-24%	-23%	-22%

The final set of IRR estimates we present, adjusts the future energy savings from a heating system replacement to correspond with the observed estimates in Figure 8. In this case, savings erode more quickly over time for households living in more deprived areas. Taking this in account results in a further reduction in the IRR for lower income households, particularly if the policy cost is taken into account.

Table 9: IRR for each measure and IMD group adjusting for time-path of energy savings

	IMD1	IMD2	IMD3	IMD4	IMD5
BoilerL	6%	12%	16%	20%	19%
BoilerU	-29%	-26%	-23%	-21%	-22%

5.3.3 Other measures of cost-effectiveness

To broaden the perspective somewhat we also consider two other measures of cost-effectiveness: the cost per tonne of CO2 removed and the cost per kWh of energy saved. These are calculated at the sample average and allow a comparison of the overall cost of these policies with other similar initiatives[5]. The estimated cost per kWh of energy saved is calculated by summing up the annual estimated savings over the expected lifetime of the measure. To calculate the cost per tonne of CO2 removed, we convert our kWh estimates based on the estimated CO2 produced in consuming one kWh of gas and electricity as calculated by the Carbon Trust (needs ref). These are reported in Table **??**, detailed information on the underlying assumptions is presented in Table F1.

Cavity wall insulation is the most cost-effective measure, followed by loft insulation and replacement heating systems. Relative to the estimated social cost of carbon and natural gas prices, insulation and the lower estimate for replacement heating systems seem relatively cost effective. However, at the upper bound of replacement heating system cost it does not represent an attractive investment.

Table 10: Cost-effectiveness of each measure

	GBP per tonne of CO2	**GBP per kWh**
Cavity wall insulation	36	0.0072
Loft insulation	90	0.0171
BoilerL	60	0.0141
BoilerU	600	0.1412

Calculated using Carbon Trust estimates of CO2 per kWh of electricity and gas

6 Discussion

This research provides an analysis of the returns to the installation of energy efficiency measures. Statistical matching and a range of panel econometric estimators are used to control for unobserved heterogeneity and selection into various government schemes which funded the upgrades. The database used covers the universe of households entering into energy efficiency schemes administered by energy suppliers in the UK.

The data allows us to examine the variation in performance depending on when measures were installed, how they perform over time, how this varies by dwelling and socioeconomic characteristics, and ultimately how this affects the cost-effectiveness of measures for different household types. The primary contribution of this work is to provide a detailed breakdown of the distributional effects of installing measures, and uniquely, we can characterise this over longer periods of time than have previously been examined.

Results indicate that cavity wall insulation and heating system replacement (installation of a condensing gas boiler) result in an energy saving of about 10 percent of annual consumption, while loft insulation results in

[5]See Appendix G for comparison with other schemes

approximately a three percent reduction. These savings are consistent regardless of when the measures were installed.

Households living in more deprived areas observe less savings (both in absolute and percentage terms) than those in more affluent areas. This result is true for all measures examined. In addition to this, savings from heating system replacements erode quickly over time for the most deprived households, but remain stable for more affluent households. As far as we are aware this result has not been shown before. Due to changes in the UK building regulations in 2005, all boiler replacements we observe are required to be condensing boilers, which are of 86% or higher efficiency. Therefore, this finding would not appear to be as a result of differences in quality. It is more likely due to a behavioural response over time, or also potentially related to variations in the frequency of servicing for different household types.

The econometric estimates are then combined with cost estimates and energy price information to examine the cost-effectiveness of each measure for different household types. Significant variation is observed, with measures having a considerably higher rate of return for households in more affluent areas, and being negative in a number of cases. Once the time-path of energy savings is considered for heating system replacement this further reduces the IRR for households in more deprived areas.

These results raise some important issues and provide new information with regard to the incentives facing households when making investments in energy efficiency, and in evaluating policies which finance efficiency improvements. If purely financial considerations are taken into account, many of these measures do not seem attractive for low-income households. However, research suggests that other welfare benefits relating to health and well-being exist. Research must do more to quantify these additional benefits.

The Green Deal was a recent policy initiative in the UK which provided households with loans in order to finance energy efficiency measures at interest rates of approximately eight percent. Given the results we observe, it is clear that this rate is not sufficiently low to provide incentives for many households to partake in this scheme. In particular, low income households would actually lose money by making these improvements unless energy prices rise significantly.

At an individual household level, the private benefits of energy efficiency investments need to be re-considered with a greater focus on the non-financial benefits. While at a societal level a greater focus on carbon emissions reduction, as opposed to cost-savings is required.

References

Adan, H. and Fuerst, F. (2015). Do energy efficiency measures really reduce household energy consumption? a difference-in-difference analysis. *Energy Efficiency*, 9(5):1207–1219.

Alberini, A. and Towe, C. (2015). Information v. energy efficiency incentives: Evidence from residential electricity consumption in maryland. *Energy Economics*, 52:S30–S40.

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics*, 130(3):1117–1165.

Allcott, H. and Greenstone, M. (2012). Is there an energy efficiency gap? *Journal of Economic Perspectives*, 26(1):3–28.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton University Press, Princeton.

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine*, 28(25):3083–3107.

BEIS (2013). Fuel poverty statistics.

Bertoldi, P. and Rezessy, S. (2008). Tradable white certificate schemes: fundamental concepts. *Energy efficiency*, 1(4):237–255.

Blackwell, M., Iacus, S. M., King, G., and Porro, G. (2009). cem: Coarsened exact matching in stata. *The Stata Journal*, 9(4):524–546.

Blumstein, C., Krieg, B., Schipper, L., and York, C. (1980). Overcoming social and institutional barriers to energy conservation. *Energy*, 5(4):355–371.

Brounen, D., Kok, N., and Quigley, J. M. (2012). Residential energy use and conservation: Economics and demographics. *European Economic Review*, 56(5):931–945.

Caliendo, M. and Kopeinig, S. (2005). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys (2008)*, 22(1):31–72.

Chawla, M., Pollitt, M. G., et al. (2013). Energy-efficiency and environmental policies & income supplements in the uk: evolution and distributional impacts on domestic energy bills. *Economics of Energy & Environmental Policy*, 2(1).

DECC (2016). Domestic NEED: Methodology Note. Technical report.

Dowson, M., Poole, A., Harrison, D., and Susman, G. (2012). Domestic uk retrofit challenge: Barriers, incentives and current performance leading into the green deal. *Energy Policy*, 50:294–305.

Duffy, A. (2013). The final report of the community energy saving programme (cesp) 2009-2012. *Ofgem, UK*.

Ferraro, P. J. and Miranda, J. J. (2017). Panel data designs and estimators as substitutes for randomized controlled trials in the evaluation of public programs. *Journal of the Association of Environmental and Resource Economists*, 4(1):281–317.

Fowlie, M., Greenstone, M., and Wolfram, C. (2015). Do energy efficiency investments deliver? evidence from the weatherization assistance program. Technical report, National Bureau of Economic Research.

Gerarden, T., Newell, R. G., and Stavins, R. N. (2015). Deconstructing the energy-efficiency gap: conceptual frameworks and evidence. *The American Economic Review*, 105(5):183–186.

Gibbons, C. E., Urbancic, M. B., Anderson, M., Auerbach, A., Andrews, R., Bitler, M., Bohn, H., Buchinsky, M., Bugni, F., Cameron, C., Dobkin, C., Khan, S., Kasy, M., Kline, P., Kodrzycki, Y., Lozano, F., Masten, M., Maurel, A., Miller, D., Montoy, J. C., Moretti, E., Oaxaca, R., Raphael, S., Rosen, A., Shapiro, J., and Sekhon, J. (2014). Broken or fixed effects? NBER Working Paper No. 20342. *NBER Working Paper No. 20342*.

Giraudet, L.-G., Bodineau, L., and Finon, D. (2012). The costs and benefits of white certificates schemes. *Energy Efficiency*, 5(2):179–199.

Golove, W. H. and Eto, J. H. (1996). Market barriers to energy efficiency: a critical reappraisal of the rationale for public policies to promote energy efficiency. *LBL-38059. Berkeley, CA: Lawrence Berkeley National Laboratory*.

Hamilton, I. G., Shipworth, D., Summerfield, A. J., Steadman, P., Oreszczyn, T., and Lowe, R. (2014). Uptake of energy efficiency interventions in english dwellings. *Building Research & Information*, 42(3):255–275.

Hamilton, I. G., Summerfield, A. J., Shipworth, D., Steadman, J. P., Oreszczyn, T., and Lowe, R. J. (2016). Energy efficiency uptake and energy savings in english houses: A cohort study. *Energy and Buildings*, 118:259–276.

Harold, J., Lyons, S., and Cullinan, J. (2015). The determinants of residential gas demand in Ireland. *Energy Economics*, 51:475–483.

Hausman, J. A. (1979). Individual discount rates and the purchase and utilization of energy-using durables. *Bell Journal of Economics*, 10(1):33–54.

Hough, D. and Page, R. (2015). Eco, the energy company obligation. *Briefing Paper: Number CBP*, 6814:1–32.

Iacus, S. M., King, G., and Porro, G. (2008). Matching for causal inference without balance checking.

Imai, K., King, G., and Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the royal statistical society: series A (statistics in society)*, 171(2):481–502.

Jaffe, A. B. and Stavins, R. N. (1994). The energy paradox and the diffusion of conservation technology. *Resource and Energy Economics*, 16(2):91–122.

Kotchen, M. J. (2017). Longer-run evidence on whether building energy codes reduce residential energy consumption. *Journal of the Association of Environmental and Resource Economists*, 4(1):135–153.

Lees, E. (2006). Evaluation of the energy efficiency commitment 2002-05. *DEFRA, Wantage (Oxfordshire).*

Lees, E. (2008). Evaluation of the energy efficiency commitment 2005-08. *DEFRA, Wantage (Oxfordshire).*

Longhi, S. (2015). Residential energy expenditures and the relevance of changes in household circumstances. *Energy Economics*, 49:440–450.

of National Statistics, O. (2011). Technical Report: Welsh Index of Multiple Deprivation 2011. Technical report.

Ofgem (2013). The final report of the Community Energy Saving Programme (CESP) 2009-2012. *Ofgem.*

Parag, Y. and Darby, S. (2009). Consumer-supplier-government triangular relations: rethinking the uk policy path for carbon emissions reduction from the uk residential sector. *Energy Policy*, 37(10):3984–3992.

Payne, R. and Abel, G. (2012). UK indices of multiple deprivation - a way to make comparisons across constituent countries easier. *Health Statistics Quarterly, ONS*, 53:22–38.

Rosenow, J. (2012). Energy savings obligations in the uka history of change. *Energy Policy*, 49:373–382.

Sorrell, S., Harrison, D., Radov, D., Klevnas, P., and Foss, A. (2009). White certificate schemes: Economic analysis and interactions with the eu ets. *Energy Policy*, 37(1):29–42.

Wooldridge, J. (2010). *Econometric analysis of cross section and panel data.* MIT Press, Cambridge, Mass., 2nd ed. edition.

7 Appendix

7.1 Balance tables for matched variables

Table A1: Unmatched sample

Unmatched sample		Treated			Control			Balance	
	Variable	Mean	Variance	Skewness	Mean	Variance	Skewness	Std-diff	Var-ratio
Variables used in matching	prop_age	2.96	1.98	0.16	3.00	3.00	0.31	-0.03	0.66
	imd_both	2.85	2.11	0.15	2.96	2.01	0.05	-0.08	1.05
	region	5.34	7.33	0.03	5.81	6.31	-0.29	-0.18	1.16
	fuel_type	0.98	0.02	-7.43	0.98	0.02	-6.33	0.04	0.74
	Gcons2005	18124	78900000	0.65	17394	86200000	0.73	0.08	0.92
Variables not used in matching	prop_type	3.33	2.63	0.22	3.56	2.92	0.08	-0.14	0.90
	floor_area	2.20	0.40	0.89	2.20	0.46	0.77	-0.01	0.86
	loft_depth	2.03	0.28	0.04	2.08	0.53	-0.13	-0.08	0.52
	wall_cons	0.73	0.20	-1.02	0.59	0.24	-0.36	0.29	0.82
	FP_ENG	2.95	1.97	0.06	2.89	2.12	0.10	0.04	0.93
	Econs2005	3903	7653561.00	2.16	3998.54	8374713	2.14	-0.03	0.91

Notes: Some text

Table A2: Matched sample all years

All years matched sample		Treated			Control			Balance	
	Variable	Mean	Variance	Skewness	Mean	Variance	Skewness	Std-diff	Var-ratio
Variables used in matching	prop_age	2.91	2.32	0.24	2.91	2.31	0.24	0.00	1.00
	imd_both	2.92	2.07	0.09	2.92	2.07	0.09	0.00	1.00
	region	5.62	6.83	-0.15	5.62	6.82	-0.15	0.00	1.00
	fuel_type	0.98	0.02	-7.45	0.98	0.02	-7.47	0.00	1.01
	Gcons2005	18020	84200000	0.67	18017	84300000	0.67	0.00	1.00
Variables not used in matching	prop_type	3.36	2.68	0.19	3.48	2.87	0.14	-0.07	0.93
	floor_area	2.21	0.42	0.86	2.21	0.45	0.80	0.00	0.93
	loft_depth	2.04	0.30	0.03	2.05	0.52	-0.08	-0.02	0.57
	wall_cons	0.67	0.22	-0.71	0.63	0.23	-0.52	0.09	0.95
	FP_ENG	2.95	2.04	0.06	2.96	2.09	0.05	-0.01	0.98
	Econs2005	3945.89	7999389.00	2.15	4028.94	8182317.00	2.09	-0.03	0.98

Notes: Some text

B Overview of Need dataset

Table B1: Description of variables in NEED

Variable	Description
HH_ID	Household identifier. A unique value for each record. Created specifically for these datasets.
REGION	Region code - formally Government Office Region. See ONS website for more details: http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/administrative/england/government-office-regions/index.html.
IMD_ENG	English Index of multiple deprivation 2010. Households are allocated to five groups (quintiles) based on the deprivation rank of the 2001 Lower Layer Super Output Area (LSOA) they are located in. Households in the 20 per cent most deprived LSOAs are in the bottom quintile (1) and households in the 20 per cent least deprived LSOAs are in the top quintile (5).
IMD_WALES	Welsh Index of multiple deprivation 2011. Households are allocated to one of five bands based on the deprivation rank of the LSOA (2001) they are located in. 1, most deprived, 5 least deprived.
FP_ENG	EUL only. Fuel Poverty Indicator for England. Households are allocated to one of five bands based on the estimate of the proportion of households in fuel poverty in the LSOA (2001) they are located in. Uses the 2011 estimates of fuel poverty low income high cost definition.
EPC_INS_DATE	EUL Only. Provides information on the date of the EPC inspection (based on lodgement date).
GconsYEAR	Weather corrected annual gas consumption. Based on meter point data from Xoserve and independent gas transporters. Readings relate to October to September each year (e.g. 2012 consumption is October 2011 to September 2012). See here for more information on this source: https://www.gov.uk/government/publications/regional-energy-data-guidance-note.
GconsYEARValid	Flag indicates records with valid consumption and households off the gas network.
EconsYEAR	Annual electricity consumption in kWh - values relate to end January to end January each year (e.g. 2012 consumption is end January 2012 to end January 2013). See here for more information on this source: https://www.gov.uk/government/publications/regional-energy-data-guidance-note.
EconsYEARValid	Valid electricity consumption (between 100 and 25,000 inclusive)
E7Flag2012	Shows whether the electricity meter is an E7 (profile 2) meter - this does not necessarily mean the household has an E7 tariff, some households will have an E7 meter without an E7 tariff.
MAIN_HEAT_FUEL	Main fuel used to heat the property, based on information from Energy Performance Certificate
PROP_AGE	Banded year of construction based on EPC data.
PROP_TYPE	Type of property (based on combination of EPC built form and property type).
FLOOR_AREA_BAND	Banded floor area based on EPC (m2).
EE_BAND	Energy Effeciency Band Based on EPC (A and B grouped).
LOFT_DEPTH	Amount of loft insulation as assessed by EPC (all properties with loft insulation recorded as installed through a Government scheme are assigned 2 irrespective of EPC information). No information could occur where the information is missing from the EPC or where the property does not have a loft.
WALL_CONS	Wall construction as recorded on EPC.
CWI	Cavity wall insulation installed through Government schemes. This includes measures recorded as installed on HEED, including, Energy Efficiency Commitment, Community Energy Savings Programme and Carbon Emissions Reduction Target.
CWI_YEAR	Year of CWI installation
LI	Loft insulation installed through Government schemes. This includes measures recorded as installed on HEED, including, Energy Efficiency Commitment, Community Energy Savings Programme and Carbon Emissions Reduction Target.
LI_YEAR	Year of LI installation
BOILER	This includes boilers installed through Government schemes, and those registered by CORGI (up to 2009) and Gas Safe (2009 onwards).
BOILER_YEAR	Year of Boiler installation
WEIGHT	EUL Only. Weighting based on Region, property age, property type and floor area band. Summing all weights gives (approximate) total number of households in England and Wales 2011.

Notes: Some text

Table B2: Summary statistic

Variable	Category	n untreated	n treated	Total n	Mean Gas	Mean Electr.
Regions	North East	83,698	115,415	199,113	15510.84	3462.747
	North West	236,794	266,874	503,668	15208.41	3900.542
	Yorkshire (a.t.H.)	188,290	201,006	389,296	15493.32	3813.363
	East Midlands	174,550	179,017	353,567	15177.16	4003.458
	West Midlands	199,366	185,984	385,350	15007.05	4146.137
	East of England	246,651	184,685	431,336	15068.45	4371.629
	London	389,548	185,003	574,551	15372.6	4135.165
	South East	363,751	270,953	634,704	15507.96	4382.642
	South West	229,357	176,074	405,431	13845.36	4457.526
	Wales	105,071	104,361	209,432	14676.42	3789.692
FP	1	515,635	359,281	874,916	13103.34	3998.862
	2	409,940	374,097	784,037	15036.81	4164.775
	3	396,893	363,103	759,996	15661.29	4219.243
	4	375,062	336,290	711,352	16214.88	4197.805
	5	414,474	332,241	746,715	16104.2	4078.469
IMD	1	450,616	455,048	905,664	12817.86	3603.685
	2	469,230	377,913	847,143	14104.08	3890.909
	3	466,882	353,377	820,259	15181.31	4246.657
	4	431,621	338,563	770,184	16325	4429.783
	5	398,727	344,471	743,198	17867.73	4495.282
Age	before 1930	606,589	391,713	998,302	16716.21	4303.14
	1930-1949	251,402	314,856	566,258	16863.97	4047.436
	1950-1966	290,718	431,623	722,341	14515.74	3873.33
	1967-1982	369,192	436,771	805,963	13831.01	3931.536
	1983-1995	299,008	201,981	500,989	13410.3	4250.797
	1996 onwards	400,167	92,428	492,595	13966.06	4287.513
Type	Detached	295,374	251,537	546,911	21843.34	5448.197
	Semi-detached	408,987	514,204	923,191	16433.14	4138.248
	End-Terrace	191,830	186,240	378,070	14938.36	3952.917
	Mid-Terrace	450,951	385,582	836,533	13771.39	3729.568
	Bungalow	138,088	238,298	376,386	15926.08	3978.159
	Flat	731,846	293,511	1,025,357	9899.633	3788.039

C Results from alternative specifications

Table C1: Summary statistic

The effect of energy efficiency upgrades on energy consumption

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
	Full sample	**Only gas**	**Matched sample**	**Only gas and matched**	**Only gas, matched and elec**	**50 drop**	**60 drop**	**70 drop**	**70 drop, elec**
Cavity wall insulation	-0.092***	-0.092***	-0.083***	-0.084***	-0.083***	-0.095***	-0.096***	-0.094***	-0.092***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Loft insulation	-0.025***	-0.026***	-0.018***	-0.019***	-0.020***	-0.029***	-0.030***	-0.030***	-0.029***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Replacement boiler	-0.055***	-0.062***	-0.038***	-0.045***	-0.049***	-0.090***	-0.092***	-0.092***	-0.091***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
					0.179***				0.138***
					(0.000)				(0.001)
Control variables	Y	Y	Y	Y	Y	Y	Y	Y	Y
Household fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year*region fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations									
Number of households									
R squared	0.115	0.118	0.115	0.167	0.118	0.398	0.375	0.349	0.369

Notes: This table reports coefficient estimates and standard errors from eight separate regressions. The dependent variable in all regressions is the logarithm of annual gas consumption in kilowatt hours. For each upgrade group a matched control group is created using coarsened-exact matching. The sample includes billing records from 2005 to 2012. Standard errors are clustered at the household level. Triple asterisks denote statistical significance at the 1% level; Double asterisks at the 1% level; single asterisks at the 10% level.

D Results in kWh

Table D1: Results in units of energy saved (kWh)

	(1) **Gas**	(2) **Elec**	(3) **Tot Energy**
Cavity wall insulation	-1254.909***	-23.002	-1277.911***
	(41.989)	-14.009	(46.216)
Loft insulation	-375.539***	-8.901	-384.440***
	(36.250)	-11.268	(39.533)
Replacement boiler	-1229.965***	-98.273***	-1328.238***
	(32.918)	-10.796	(35.722)
Control variables	Y	Y	Y
Household fixed effects	Y	Y	Y
Year fixed effects	Y	Y	Y
Year*region fixed effects	Y	Y	Y
Observations	549072	549072	549072
Number of households	68,634	68634	68634
R squared	0.337	0.092	0.357

Notes: This table reports coefficient estimates and standard errors from eight separate regressions. The dependent variable in all regressions is the logarithm of annual gas consumption in kilowatt hours. Column(1) "All" denotes efficiency upgrades occuring at any time during the sample period. Columns (2-8) relate to upgrades occuring only in the relevant year. Each individual year denotes upgrades occuring solely in that year. For each upgrade group a matched control group is created using coarsened-exact matching. The sample includes billing records from 2005 to 2012. Standard errors are clustered at the household level. Triple asterisks denote statistical significance at the 1% level; Double asterisks at the 1% level; single asterisks at the 10% level.

E Cost assumptions

Table E1: Sources of cost assumptions

	Low ()	High ()
Cavity wall (pre 1976)	300	325
Cavity wall (post 1976)	300	325
Loft 300mm (currently none)	138	273
Loft 300mm (currently 100mm)	86	211
Loft 300mm (currently 200mm)	35	170
Condensing boiler	100	300

Notes: Source Shorrock (2005).

Table E2: Sources of cost assumptions

	EESOP1 (1994)	EESOP2	EESOP3	EEC1 (2005)
Cavity wall insulation	223	219	261	261
Condensing boiler	450	270	165	114

Notes: Source Lees (2005).

Table E3: Sources of cost assumptions

	Defra EEC1	Defra EEC2	Defra CERT	Lees 2005	Lees 2008
Cavity wall insulation	268	313	380	274	350
Loft insulation (top up)	213	260	286	217	275
Loft insulation (virgin)	213	260	286	252	295
A and B boiler	145			120	
A and B boiler and heating control	217			190	
All boilers			50		45

Notes: Source Lees (2005, 2008).

F Carbon saving assumptions

Table F1: Assumptions for CO2 saving calculation

	(1) Cavity wall insulation		(2) Loft insulation		(3) Replacement heating system	
	Gas	**Elec**	**Gas**	**Elec**	**Gas**	**Elec**
kg CO2 per (kWh)	0.18	0.52	0.18	0.52	0.18	0.52
Total annual saving (kWh)	1551.46	67.51	546.22	10.74	1363.85	52.29
Total annual saving (kgCO2)	284.85	35.42	100.29	5.63	250.40	27.43
Lifespan	30	30	30	30	12	12
Total lifetime savings (kgCO2)	8545.41	1062.54	3008.56	168.99	3004.82	329.16

Notes: Some text

G Comparison of cost-effectiveness

Table H1: Cost assumptions for each measure

Intervention type	Reference	Evaluation type	Relevant subset	Percent reduction in energy usage	Engineering estimates of percent reduction in energy usage	Cost effectiveness (cents per kWh saved, 2015 USD)
Behavioral programs	Allcott (2011)	RCT	NA	2		3.6
	Allcott & Rogers (2014)	RCT	One-shot intervention			4.4
			Two-year intervention			1.1 to 1.8
			Four-year intervention			1.2 to 1.8
	Ayres et al. (2012)	RCT	Sacramento, California	2		5.5
			Puget Sound, Washington	1.2		2
Building codes	Novan et al. (2017)c	RD analysis	NA	1.3	20	24.4
Efficient equipment or energy savings subsidy	Alberini & Towe (2015)	Matching	NA	5.3		3.9
	Alberini et al. (2016)	DID	Rebate of $1,000 or more	0		
			Rebate of $450	5.5		47.9
			Rebate of $300	6.2		28.2
	Burlig et al. (2017)	Machine learning	NA	2.9 to 4.5	11.6 to 18	
	Davis et al. (2014)	DID regression	Refrigerators	8		27.2
			Air conditioners	plus 1.7		4.5
Information provision	Alberini & Towe (2015)	Matching		5.5		
Supplier Obligation (TWC)	McCoy & Kotsch (2018)	Matching, FE regression	Cavity wall insulation	9.4	20.0	1.54 to 2.31
			Loft insulation	3	5.2	3.65 to 5.47
			Replacement heating system	9.2	24.9	3.02 to 30.19
Previous estimate of UK Supplier obligation Lees, 2008)						1.92

Adapted from Gillingham et al (2018)